

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
Please do not report the images to the
Image Problem Mailbox.

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-319396
 (43)Date of publication of application : 12.12.1997

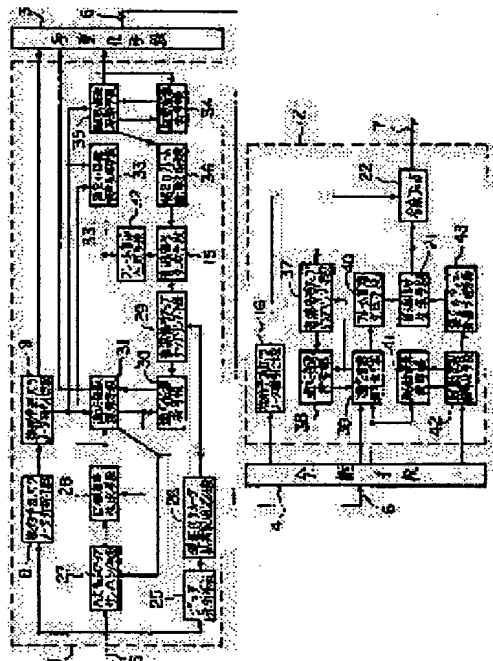
1017 U.S. PTO
 10/072892
 02/12/02

(51)Int.Cl. G10L 9/14
 G10L 9/18
 H03M 7/30

(21)Application number : 08-135240 (71)Applicant : MITSUBISHI ELECTRIC CORP
 (22)Date of filing : 29.05.1996 (72)Inventor : YAMAURA TADASHI
 TAZAKI HIROHISA
 TAKAHASHI MASAYA

(54) SPEECH ENCODING DEVICE, AND SPEECH ENCODING AND DECODING DEVICE**(57)Abstract:**

PROBLEM TO BE SOLVED: To avert the deterioration of the quality of synthesized speech and to form a synthesized speech having good quality with smaller quantity of computation at the time of encoding speech by evaluating the distortion of the synthesized speech vector obtd. from an adaptive sound source vector with respect to a target speech vector and searching the adaptive sound source vector to minimize the distortion. **SOLUTION:** This device has a target speech forming means 28, an adaptive section code book 30, an adaptive sound source searching means 31 and a frame sound source forming means 32. If a delay parameter is shorter than a frame length at the time of determining an optimum delay parameter, the target speech vector of the vector length is formed by periodically adding and averaging input speeches. The distortion from the synthesized speech vector formed by a linear prediction synthesis of the adaptive sound source vector of the vector length for the same is evaluated. The synthesized



speech vector formed by the linear prediction synthesis of the driving sound source vector of the vector length is used for distortion evaluation at the time of determining optimal driving sound source codes. Then, the deterioration in the quality of the synthesized speeches is averted.

LEGAL STATUS

[Date of request for examination] 23.03.2000
 [Date of sending the examiner's decision of rejection]
 [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
 [Date of final disposal for application]

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/14			G 1 0 L 9/14	G
	9/18		9/18	E
H 0 3 M 7/30		9382-5K	H 0 3 M 7/30	A

審査請求 未請求 請求項の数23 O L (全 21 頁)

(21) 出願番号 特願平8-135240

(22) 出願日 平成8年(1996)5月29日

(71) 出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72) 発明者 山浦 正

東京都千代田区丸の内二丁目2番3号 三

菱電機株式会社内

(72) 発明者 田崎 裕久

東京都千代田区丸の内二丁目2番3号 三

菱電機株式会社内

(72) 発明者 高橋 真哉

東京都千代田区丸の内二丁目2番3号 三

菱電機株式会社内

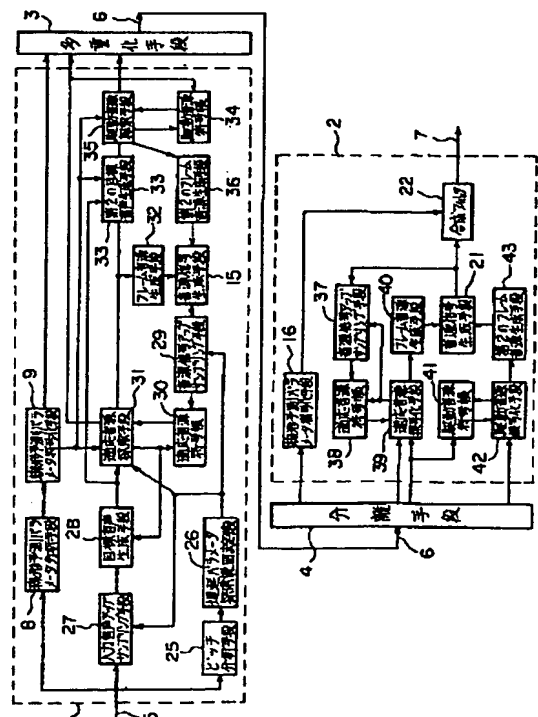
(74) 代理人 井理士 宮田 金雄 (外3名)

(54) 【発明の名称】 音声符号化装置および音声符号化復号化装置

(57) 【要約】

【課題】 音声を符号化するに当たり、合成音声の品質の劣化を回避し、少ない演算量で品質の良い合成音声を生成することができる音声符号化装置を得る。

【解決手段】 音声符号化装置は、入力音声から遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段と、過去に生成した音源信号から前記遅延パラメータに対応したベクトル長の適応音源ベクトルを生成する適応音源符号帳と、前記適応音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる適応音源ベクトルを探索する適応音源探索手段と、前記歪みが最小となる適応音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備える。



【特許請求の範囲】

【請求項 1】 入力音声をスペクトル包絡情報と音源信号情報に分けて、フレーム単位に音源信号情報を符号化する音声符号化装置において、入力音声から遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段と、過去に生成した音源信号から前記遅延パラメータに対応したベクトル長の適応音源ベクトルを生成する適応音源符号帳と、前記適応音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる適応音源ベクトルを探索する適応音源探索手段と、前記歪みが最小となる適応音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えたことを特徴とする音声符号化装置。

【請求項 2】 目標音声ベクトルと歪みが最小となる適応音源ベクトルから第 2 の目標音声ベクトルを生成する第 2 の目標音声生成手段と、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、前記駆動音源ベクトルから得られる第 2 の合成音声ベクトルの前記第 2 の目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段と、前記歪みが最小となる駆動音源ベクトルから第 2 のフレーム長の音源信号を生成する第 2 のフレーム音源生成手段とを備えたことを特徴とする請求項 1 記載の音声符号化装置。

【請求項 3】 入力音声をスペクトル包絡情報と音源信号情報に分けて、フレーム単位に音源信号情報を符号化する音声符号化装置において、入力音声から遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段と、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、前記駆動音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段と、前記歪みが最小となる駆動音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えたことを特徴とする音声符号化装置。

【請求項 4】 遅延パラメータは入力音声のピッチ周期に対応して決定することを特徴とする請求項 3 記載の音声符号化装置。

【請求項 5】 遅延パラメータに対応したベクトル長が有理数をとることを特徴とする請求項 1～請求項 3 のいずれかに記載の音声符号化装置。

【請求項 6】 目標音声生成手段は、フレームの入力音声を遅延パラメータに対応したベクトル長毎に分割し、ベクトル長毎の入力音声を加重平均して目標音声ベクトルを生成することを特徴とする請求項 1～請求項 3 のいずれかに記載の音声符号化装置。

【請求項 7】 目標音声生成手段は、遅延パラメータに対応したベクトル長の整数倍長の入力音声をベクトル長

毎に分割し、ベクトル長毎の入力音声を加重平均して目標音声ベクトルを生成することを特徴とする請求項 1～請求項 3 のいずれかに記載の音声符号化装置。

【請求項 8】 遅延パラメータに対応したベクトル長の整数倍長はフレーム長以上とすることを特徴とする請求項 7 記載の音声符号化装置。

【請求項 9】 目標音声生成手段は、遅延パラメータに対応したベクトル長毎の入力音声に関わる特徴量に応じて、入力音声をベクトル長毎に加重平均して目標音声ベクトルを生成する際の重みを決定することを特徴とする請求項 6 又は請求項 7 記載の音声符号化装置。

【請求項 10】 遅延パラメータに対応したベクトル長毎の入力音声に関わる特徴量は少なくとも入力音声のパワー情報を含むことを特徴とする請求項 9 記載の音声符号化装置。

【請求項 11】 遅延パラメータに対応したベクトル長毎の入力音声に関わる特徴量は少なくとも入力音声の相関情報を含むことを特徴とする請求項 9 に記載の音声符号化装置。

【請求項 12】 目標音声生成手段は、遅延パラメータに対応したベクトル長毎の入力音声の時間関係に応じて、入力音声をベクトル長毎に加重平均して目標音声ベクトルを生成する際の重みを決定することを特徴とする請求項 6 又は請求項 7 記載の音声符号化装置。

【請求項 13】 目標音声生成手段は、入力音声を遅延パラメータに対応したベクトル長毎に加重平均する際、ベクトル長毎の入力音声の時間関係を微調整することを特徴とする請求項 6 又は請求項 7 記載の音声符号化装置。

【請求項 14】 フレーム音源生成手段は、遅延パラメータに対応したベクトル長の音源ベクトルをベクトル長毎に繰り返して周期化し、フレーム長の音源信号を生成することを特徴とする請求項 1～請求項 3 のいずれかに記載の音声符号化装置。

【請求項 15】 フレーム音源生成手段は、遅延パラメータに対応したベクトル長の音源ベクトルをフレーム間で補間して音源信号を生成することを特徴とする請求項 1～請求項 3 のいずれかに記載の音声符号化装置。

【請求項 16】 適応音源探索手段は合成フィルタを備え、この合成フィルタのインパルス応答を用いて、適応音源ベクトルから得られる合成音声ベクトルの目標音声ベクトルに対する歪みを反復的に計算することを特徴とする請求項 1 記載の音声符号化装置。

【請求項 17】 入力音声をアップサンプリングする入力音声アップサンプリング手段を備え、目標音声生成手段はアップサンプリングされた入力音声から目標音声ベクトルを生成することを特徴とする請求項 5 記載の音声符号化装置。

【請求項 18】 過去に生成された音源信号をアップサンプリングする音源信号アップサンプリング手段を備

え、適応音源符号帳はアップサンプリングされた過去に生成された音源信号から適応音源ベクトルを生成することを特徴とする請求項 5 記載の音声符号化装置。

【請求項 19】 アップサンプリング手段は遅延パラメータに対応してアップサンプリング倍率を変更することを特徴とする請求項 17 又は請求項 18 記載の音声符号化装置。

【請求項 20】 アップサンプリング手段は、遅延パラメータに対応したベクトル長に応じた範囲のみ入力音声または音源信号のアップサンプリング倍率を変更することを特徴とする請求項 17 又は請求項 18 記載の音声符号化装置。

【請求項 21】 入力音声をスペクトル包絡情報と音源信号情報に分けて、フレーム単位に音源信号情報を符号化し、符号化された音源信号情報を復号化して出力音声を生成する音声符号化復号化装置において、符号化側に、入力音声から遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段と、過去に生成した音源信号から前記遅延パラメータに対応したベクトル長の適応音源ベクトルを生成する適応音源符号帳と、前記適応音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる適応音源ベクトルを探索する適応音源探索手段と、前記歪みが最小となる適応音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備える一方、復号化側に、遅延パラメータに対応したベクトル長の適応音源ベクトルを生成する適応音源符号帳と、適応音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えることを特徴とする音声符号化復号化装置。

【請求項 22】 符号化側に、目標音声ベクトルと適応音源ベクトルから第 2 の目標音声ベクトルを生成する第 2 の目標音声生成手段と、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、前記駆動音源ベクトルから得られる第 2 の合成音声ベクトルの前記第 2 の目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段と、前記歪みが最小となる駆動音源ベクトルから第 2 のフレーム長の音源信号を生成する第 2 のフレーム音源生成手段とを備える一方、復号化側に、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、駆動音源ベクトルから第 2 のフレーム長の音源信号を生成する第 2 のフレーム音源生成手段とを備えることを特徴とする請求項 21 記載の音声符号化復号化装置。

【請求項 23】 入力音声をスペクトル包絡情報と音源信号情報に分けて、フレーム単位に音源信号情報を符号化し、符号化された音源信号情報を復号化して出力音声を生成する音声符号化復号化装置において、符号化側に、入力音声から遅延パラメータに対応したベクトル長

の目標音声ベクトルを生成する目標音声生成手段と、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、前記駆動音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段と、前記歪みが最小となる駆動音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備える一方、復号化側に、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、駆動音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えることを特徴とする音声符号化復号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、音声信号をディジタル信号に圧縮符号化する音声符号化装置および音声符号化復号化装置に関する。

【0002】

【従来の技術】図 9 は、従来の入力音声をスペクトル包絡情報と音源信号情報に分けて、フレーム単位に音源信号情報を符号化する音声符号化復号化装置の全体構成の一例を示すものであり、特開昭 64-40899 号公報に示されたものと同様のものである。

【0003】図において、1 は符号化部、2 は復号化部、3 は多重化手段、4 は分離手段、5 は入力音声、6 は伝送路、7 は出力音声である。符号化部 1 は次の 8~15 により構成されている。8 は線形予測パラメータ分析手段、9 は線形予測パラメータ符号化手段、10 は適応音源符号帳、11 は適応音源探索手段、12 は誤差信号生成手段、13 は駆動音源符号帳、14 は駆動音源探索手段、15 は音源信号生成手段である。また、復号化部 2 は次の 16~22 により構成されている。16 は線形予測パラメータ復号化手段、17 は適応音源符号帳、18 は適応音源復号化手段、19 は駆動音源符号帳、20 は駆動音源復号化手段、21 は音源信号生成手段、22 は合成フィルタである。

【0004】以下、上記従来の入力音声をスペクトル包絡情報と音源信号情報に分けて、フレーム単位に音源信号情報を符号化する音声符号化復号化装置の動作について説明する。まず符号化部 1 において、例えば 8 kHz でサンプリングされたディジタル音声信号が入力音声 5 として入力される。線形予測パラメータ分析手段 8 は、入力音声 5 を分析して、音声のスペクトル包絡情報である線形予測パラメータを抽出する。次いで線形予測パラメータ符号化手段 9 が抽出した前記線形予測パラメータを量子化し、それに対応する符号を多重化手段 3 に出力するとともに、量子化した線形予測パラメータを適応音源探索手段 11、誤差信号生成手段 12、駆動音源探索手段 14 に出力する。

【0005】次に、音源信号情報の符号化について説明

する。適応音源符号帳 10 には音源信号生成手段 15 から入力される過去に生成した音源信号が記憶されており、適応音源探索手段 11 より入力される遅延パラメータ L に対応したフレーム長の適応音源ベクトルを適応音源探索手段 11 に出力する。ここで、前記適応音源ベクトルは遅延パラメータ L に対して 1 サンプル過去からフレーム長の音源信号を切り出したものであり、 L がフレーム長より短い場合は 1 サンプルの音源信号をフレーム長になるまで繰り返して生成したものである。図 10

(a) に $L \geq$ フレーム長の場合、図 10 (b) に $L <$ フレーム長の場合の適応音源ベクトルの例を示す。

【0006】適応音源探索手段 11 は、例えば $20 \leq L \leq 128$ の範囲の遅延パラメータ L に対して、前記適応音源符号帳 10 から入力される適応音源ベクトルを前記線形予測パラメータ符号化手段 9 から入力された量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、入力音声 5 からフレーム毎に切り出した入力音声ベクトルと前記合成音声ベクトルとの聴覚重み付き歪みを求める。次に、前記歪みを比較評価し、前記歪みが最小になる遅延パラメータ L とそれに対応する適応音源利得 β を求め、前記遅延パラメータ L と適応音源利得 β の符号を多重化手段 3 に出力するとともに、前記遅延パラメータ L に対応する適応音源ベクトルに前記適応音源利得 β を乗じた適応音源信号を生成し、誤差信号生成手段 12 と音源信号生成手段 15 に出力する。

【0007】誤差信号生成手段 12 は前記適応音源探索手段 11 から入力された適応音源信号を前記線形予測パラメータ符号化手段 9 から入力された量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、入力音声 5 からフレーム毎に切り出した入力音声ベクトルと前記合成音声ベクトルとの差分である誤差信号ベクトルを求め、駆動音源探索手段 14 に出力する。

【0008】駆動音源符号帳 13 には、例えばランダム雑音から生成した N 個の駆動音源ベクトルが記憶されており、駆動音源探索手段 14 より入力される駆動音源符号 i に対応した駆動音源ベクトルを出力する。駆動音源探索手段 14 は N 個の駆動音源ベクトルに対して、前記駆動音源符号帳 13 から入力される駆動音源ベクトルを前記線形予測パラメータ符号化手段 9 から入力された量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、前記誤差信号生成手段 12 から入力された誤差信号ベクトルと前記合成音声ベクトルとの聴覚重み付き歪みを求める。次に、前記歪みを比較評価し、前記歪みが最小になる駆動音源符号 I とそれに対応する駆動音源利得 γ を求め、前記駆動音源符号 I と駆動音源利得 γ の符号を多重化手段 3 に出力するとともに、前記駆動音源符号 I に対応する駆動音源ベクトルに前記駆動音源利得 γ を乗じた駆動音源信号

を生成し、音源信号生成手段 15 に出力する。

【0009】音源信号生成手段 15 は前記適応音源探索手段 11 から入力された適応音源信号と前記駆動音源探索手段 14 から入力された駆動音源信号とを加算して音源信号を生成し、適応音源符号帳 10 に出力する。以上符号化が終了した後、多重化手段 3 は前記量子化した線形予測パラメータに対応する符号、遅延パラメータ L 、駆動音源符号 I 、及び音源利得 β 、 γ に対応する符号を伝送路 6 に送出する。

【0010】次に、復号化部 2 の動作について説明する。まず多重化手段 3 の出力を受けた分離手段 4 は、線形予測パラメータの符号 \rightarrow 線形予測パラメータ復号化手段 16

遅延パラメータ L 、音源利得 β の符号 \rightarrow 適応音源復号化手段 18

駆動音源符号 I 、音源利得 γ の符号 \rightarrow 駆動音源復号化手段 20

にそれぞれ出力する。

【0011】線形予測パラメータ復号化手段 16 は、前記線形予測パラメータの符号に対応する線形予測パラメータを復号化し、合成フィルタ 22 に出力する。適応音源復号化手段 18 は、前記遅延パラメータ L に対応した適応音源ベクトルを適応音源符号帳 17 から読み出し、また前記適応音源利得 β の符号から適応音源利得 β を復号化し、前記適応音源ベクトルに前記適応音源利得 β を乗じた適応音源信号を生成し、音源信号生成手段 21 に出力する。駆動音源復号化手段 20 は、前記駆動音源符号 I に対応した駆動音源ベクトルを駆動音源符号帳 19 から読み出し、また前記駆動音源利得 γ の符号から駆動音源利得 γ を復号化し、前記駆動音源ベクトルに前記駆動音源利得 γ を乗じた駆動音源信号を生成し、音源信号生成手段 21 に出力する。

【0012】音源信号生成手段 21 は前記適応音源復号化手段 18 から入力された適応音源信号と前記駆動音源復号化手段 20 から入力された駆動音源信号とを加算して音源信号を生成し、適応音源符号帳 17 と合成フィルタ 22 に出力する。合成フィルタ 22 は前記音源信号生成手段 21 から入力された音源信号を、前記線形予測パラメータ復号化手段 16 から入力された線形予測パラメータを用いて線形予測合成し、出力音声 7 を出力する。

【0013】また、上記従来の音声符号化復号化装置の改良された先行技術として、より高品質な出力音声を得られる音声符号化復号化装置として、P. Kroon and B. S. Atal 著 "Pitch predictors with high temporal resolution" (ICASSP'90, pp 661-664, 1990) に示されたものがある。

【0014】この従来の改良された音声符号化復号化装置は、図 9 に示す従来の音声符号化復号化装置の構成において、適応音源探索手段 11 における探索対象の遅延

パラメータとして整数値の他に非整数有理数もとるようにし、適応音源符号帳 10、17 は前記非整数有理数の遅延パラメータに対応した適応音源ベクトルを過去に生成した音源信号のサンプル間を補間して生成し、出力するものである。図 11 に遅延パラメータ l が非整数有理数の場合の適応音源ベクトルの例を示す。図 11 (a) は $l \geq$ フレーム長の場合、図 11 (b) は $l <$ フレーム長の場合の例である。このように構成することにより、入力音声のサンプリング周期よりも高い精度で遅延パラメータを決定し、適応音源ベクトルを生成することができ、特開昭 64-40899 号公報に開示されている装置に比べてより高品質な出力音声の生成が可能である。

【0015】また、従来の音声符号化復号化装置の他の先行技術として、特開平 4-344699 号公報がある。図 12 は、この従来の音声符号化装置の全体構成の一例を示す構成図である。図 12 において図 9 と同一の部分については同一の符号を付し、その説明は省略する。図 12 において、23、24 は駆動音源符号帳で、図 9 の駆動音源符号帳とは異なるものである。

【0016】ここで上記の構成による符号化復号化装置の動作について説明する。まず、符号化部 1 において、適応音源探索手段 11 は、例えば $20 \leq 1 \leq 128$ の範囲の遅延パラメータ l に対して、適応音源符号帳 10 から入力される適応音源ベクトルを線形予測パラメータ符号化手段 9 から入力される量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、入力音声 5 からフレーム毎に切り出した入力音声ベクトルと前記合成音声ベクトルとの聴覚重み付き歪みを求める。次に、前記歪みを比較評価し、前記歪みが最小になる遅延パラメータ L とそれに対応する適応音源利得 β を求め、前記遅延パラメータ L と適応音源利得 β の符号を多重化手段 3 と駆動音源符号帳 23 に出力するとともに、前記遅延パラメータ L に対応する適応音源ベクトルに前記適応音源利得 β を乗じた適応音源信号を生成し、誤差信号生成手段 12 と音源信号生成手段 15 に出力する。

【0017】駆動音源符号帳 23 には、例えばランダム雑音から生成した N 個の駆動音源ベクトルが記憶されており、駆動音源探索手段 14 より入力される駆動音源符号 i に対応した駆動音源ベクトルを前記遅延パラメータ L に対応した周期毎に繰り返して周期化し、出力する。図 13 (a) に周期化した駆動音源ベクトルの例を示す。遅延パラメータ L が非整数有理数の場合は、図 13 (b) に示すように駆動音源ベクトルのサンプル間を補間して生成し、周期化する。

【0018】駆動音源探索手段 14 は N 個の駆動音源ベクトルに対して、前記駆動音源符号帳 23 から入力される周期化した駆動音源ベクトルを線形予測パラメータ符号化手段 9 から入力される量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成す

る。そして、誤差信号生成手段 12 から入力される誤差信号ベクトルと前記合成音声ベクトルとの聴覚重み付き歪みを求める。次に、前記歪みを比較評価し、前記歪みが最小になる駆動音源符号 I とそれに対応する駆動音源利得 γ を求め、前記駆動音源符号 I と駆動音源利得 γ の符号を多重化手段 3 に出力するとともに、前記駆動音源符号 I に対応する周期化した駆動音源ベクトルに前記駆動音源利得 γ を乗じた駆動音源信号を生成し、音源信号生成手段 15 に出力する。

【0019】符号化が終了した後、多重化手段 3 は前記量子化した線形予測パラメータに対応する符号、遅延パラメータ L 、駆動音源符号 I 、及び音源利得 β 、 γ に対応する符号を伝送路 6 に送出する。

【0020】次に、復号化部 2 の動作について説明する。まず、多重化手段 3 の出力を受けた分離手段 4 は、線形予測パラメータの符号→線形予測パラメータ復号化手段 16

遅延パラメータ L 、音源利得 β の符号→適応音源復号化手段 18、駆動音源符号帳 24

駆動音源符号 I 、音源利得 γ の符号→駆動音源復号化手段 20

にそれぞれ出力する。

【0021】駆動音源符号帳 24 は符号化側の駆動音源符号帳 23 と同じ N 個の駆動音源ベクトルが記憶されており、駆動音源復号化手段 20 より入力される駆動音源符号 I に対応した駆動音源ベクトルを前記遅延パラメータ L に対応した周期毎に繰り返して周期化し、駆動音源復号化手段 20 に出力する。

【0022】駆動音源復号化手段 20 は、前記駆動音源利得 γ の符号から駆動音源利得 γ を復号化し、前記駆動音源符号帳 24 から入力された周期化した駆動音源ベクトルに前記駆動音源利得 γ を乗じた駆動音源信号を生成し、音源信号生成手段 21 に出力する。

【0023】音源信号生成手段 21 は適応音源復号化手段 18 から入力される適応音源信号と前記駆動音源復号化手段 20 から入力された駆動音源信号とを加算して音源信号を生成し、適応音源符号帳 17 と合成フィルタ 22 に出力する。合成フィルタ 22 は前記音源信号生成手段 21 から入力された音源信号を、線形予測パラメータ復号化手段 16 から入力される線形予測パラメータを用いて線形予測合成し、出力音声 7 を出力する。

【0024】

【発明が解決しようとする課題】上記した従来の音声符号化復号化装置では、符号化における音源探索に際し、遅延パラメータに応じて、適応音源ベクトルまたは駆動音源ベクトルを周期化、生成してフレーム長の音源ベクトルとし、これを線形予測合成して合成音声ベクトルを生成し、フレーム長区間での入力音声ベクトルと合成音声ベクトルとの歪みを求めている。しかし、線形予測合成に掛かる演算量が大きいため、音源探索には多大な演

算量を要するという問題があった。

【0025】この発明は、かかる課題を解決するためになされたもので、その目的は、音声を符号化するに当たり、合成音声の品質の劣化を回避し、少ない演算量で品質の良い合成音声を生成することができる音声符号化装置及び音声符号化復号化装置を得るものである。

【0026】

【課題を解決するための手段】上記の課題を解決するために、この発明の音声符号化装置は、入力音声から遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段と、過去に生成した音源信号から前記遅延パラメータに対応したベクトル長の適応音源ベクトルを生成する適応音源符号帳と、前記適応音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる適応音源ベクトルを探索する適応音源探索手段と、前記歪みが最小となる適応音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えたものである。

【0027】また、この発明の音声符号化装置はさらに、目標音声ベクトルと歪みが最小となる適応音源ベクトルから第2の目標音声ベクトルを生成する第2の目標音声生成手段と、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、前記駆動音源ベクトルから得られる第2の合成音声ベクトルの前記第2の目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段と、前記歪みが最小となる駆動音源ベクトルから第2のフレーム長の音源信号を生成する第2のフレーム音源生成手段とを備えたものである。

【0028】また、この発明の音声符号化装置は、入力音声から遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段と、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、前記駆動音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段と、前記歪みが最小となる駆動音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えたものである。

【0029】また、この発明の音声符号化装置はさらに、目標音声ベクトル及び駆動音源ベクトルのベクトル長は入力音声のピッチ周期に対応して決定するものである。

【0030】また、この発明の音声符号化装置はさらに、遅延パラメータに対応したベクトル長が有理数をとるものである。

【0031】また、この発明の音声符号化装置はさらに、目標音声生成手段は、フレームの入力音声を遅延パラメータに対応したベクトル長毎に分割し、ベクトル長

毎の入力音声を加重平均して目標音声ベクトルを生成するものである。

【0032】また、この発明の音声符号化装置はさらに、目標音声生成手段は、遅延パラメータに対応したベクトル長の整数倍長の入力音声をベクトル長毎に分割し、ベクトル長毎の入力音声を加重平均して目標音声ベクトルを生成するものである。

【0033】また、この発明の音声符号化装置はさらに、遅延パラメータに対応したベクトル長の整数倍長はフレーム長以上とするものである。

【0034】また、この発明の音声符号化装置はさらに、目標音声生成手段は遅延パラメータに対応したベクトル長毎の入力音声に関わる特徴量に応じて、入力音声をベクトル長毎に加重平均して目標音声ベクトルを生成する際の重みを決定するものである。

【0035】また、この発明の音声符号化装置はさらに、遅延パラメータに対応したベクトル長毎の入力音声に関わる特徴量は少なくとも入力音声のパワー情報を含むものである。

【0036】また、この発明の音声符号化装置はさらに、遅延パラメータに対応したベクトル長毎の入力音声に関わる特徴量は少なくとも入力音声の相関情報を含むものである。

【0037】また、この発明の音声符号化装置はさらに、目標音声生成手段は、遅延パラメータに対応したベクトル長毎の入力音声の時間関係に応じて、入力音声をベクトル長毎に加重平均して目標音声ベクトルを生成する際の重みを決定するものである。

【0038】また、この発明の音声符号化装置はさらに、目標音声生成手段は、入力音声を遅延パラメータに対応したベクトル長毎に加重平均する際、ベクトル長毎の入力音声の時間関係を微調整するものである。

【0039】また、この発明の音声符号化装置はさらに、フレーム音源生成手段は、遅延パラメータに対応したベクトル長の音源ベクトルを前記ベクトル長毎に繰り返して周期化し、フレーム長の音源信号を生成するものである。

【0040】また、この発明の音声符号化装置はさらに、フレーム音源生成手段は、遅延パラメータに対応したベクトル長の音源ベクトルをフレーム間で補間して音源信号を生成するものである。

【0041】また、この発明の音声符号化装置はさらに、適応音源探索手段は合成フィルタを備え、この合成フィルタのインパルス応答を用いて、適応音源ベクトルから得られる合成音声ベクトルの目標音声ベクトルに対する歪みを反復的に計算するものである。

【0042】また、この発明の音声符号化装置はさらに、入力音声をアップサンプリングする入力音声アップサンプリング手段を備え、目標音声生成手段はアップサンプリングされた入力音声から目標音声ベクトルを生成

するものである。

【0043】また、この発明の音声符号化装置はさらに、過去に生成された音源信号をアップサンプリングする音源信号アップサンプリング手段を備え、適応音源符号帳はアップサンプリングされた過去に生成された音源信号から適応音源ベクトルを生成するものである。

【0044】また、この発明の音声符号化装置はさらに、アップサンプリング手段は遅延パラメータに対応して、アップサンプリング倍率を変更するものである。

【0045】また、この発明の音声符号化装置はさらに、アップサンプリング手段は遅延パラメータに対応したベクトル長に応じた範囲のみ入力音声または音源信号のアップサンプリング倍率を変更するものである。

【0046】また、この発明に係る音声符号化復号化装置は、符号化側に、入力音声から遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段と、過去に生成した音源信号から前記遅延パラメータに対応したベクトル長の適応音源ベクトルを生成する適応音源符号帳と、前記適応音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる適応音源ベクトルを探索する適応音源探索手段と、前記歪みが最小となる適応音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備える一方、復号化側に、遅延パラメータに対応したベクトル長の適応音源を生成する適応音源符号帳と、適応音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えるものである。

【0047】また、この発明に係る音声符号化復号化装置はさらに、符号化側に、目標音声ベクトルと適応音源ベクトルから第2の目標音声ベクトルを生成する第2の目標音声生成手段と、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、前記駆動音源ベクトルから得られる第2の合成音声ベクトルの前記第2の目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段と、前記歪みが最小となる駆動音源ベクトルから第2のフレーム長の音源信号を生成する第2のフレーム音源生成手段とを備える一方、復号化側に遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、駆動音源ベクトルから第2のフレーム長の音源信号を生成する第2のフレーム音源生成手段とを備えるものである。

【0048】また、この発明に係る音声符号化復号化装置は、符号化側に、入力音声からベクトル長の目標音声ベクトルを生成する目標音声生成手段と、ベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、前記駆動音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段と、前

記歪みが最小となる駆動音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備える一方、復号化側に、遅延パラメータに対応したベクトル長の駆動音源ベクトルを生成する駆動音源符号帳と、駆動音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えるものである。

【0049】

【発明の実施の形態】

実施の形態1. 図1は、この発明の実施の形態1である音声符号化装置及び音声復号化装置の全体構成を示すブロック図である。

【0050】図1において、1は符号化部、2は復号化部、3は多重化手段、4は分離手段、5は入力音声、6は伝送路、7は出力音声である。

【0051】符号化部1は次の8、9、15、および25～36により構成されている。8は線形予測パラメータ分析手段、9は線形予測パラメータ符号化手段、15は音源信号生成手段、25は入力音声のピッチ周期を抽出するピッチ分析手段、26は適応音源ベクトルを探索する際の遅延パラメータの探索範囲を決定する遅延パラメータ探索範囲決定手段、27は入力音声をアップサンプリングする入力音声アップサンプリング手段、28は遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段、29は過去に生成した音源信号をアップサンプリングする音源信号アップサンプリング手段、30は過去に生成した音源信号から遅延パラメータに対応したベクトル長の適応音源ベクトルを出力する適応音源符号帳、31は適応音源ベクトルから得られる合成音声ベクトルの目標音声ベクトルに対する歪みを評価し、歪みが最小となる適応音源ベクトルを探索する適応音源探索手段、32は遅延パラメータに対応したベクトル長の適応音源信号からフレーム長の適応音源信号を生成するフレーム音源生成手段、33は駆動音源ベクトル探索における遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する第2の目標音声生成手段、34は遅延パラメータに対応したベクトル長の駆動音源ベクトルを出力する駆動音源符号帳、35は駆動音源ベクトルから得られる合成音声ベクトルの第2の目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段、36は遅延パラメータに対応したベクトル長の駆動音源信号からフレーム長の駆動音源信号を生成する第2のフレーム音源生成手段である。

【0052】また、復号化部2は次の16、21、22、および37～43により構成されている。16は線形予測パラメータ復号化手段、21は音源信号生成手段、22は合成フィルタ、37は過去に生成した音源信号をアップサンプリングする音源信号アップサンプリング手段、38は遅延パラメータに対応したベクトル長の適応音源ベクトルを出力する適応音源符号帳、39は遅

延パラメータに対応したベクトル長の適応音源信号を復号化する適応音源復号化手段、40は遅延パラメータに対応したベクトル長の適応音源信号からフレーム長の適応音源信号を生成するフレーム音源生成手段、41は遅延パラメータに対応したベクトル長の駆動音源ベクトルを出力する駆動音源符号帳、42は遅延パラメータに対応したベクトル長の駆動音源信号を復号化する駆動音源復号化手段、43は遅延パラメータに対応したベクトル長の駆動音源信号からフレーム長の駆動音源信号を生成する第2のフレーム音源生成手段である。

【0053】以下、動作を説明する。まず符号化部1において、例えば8kHzでサンプリングされたデジタル音声信号が入力音声5として入力される。線形予測パラメータ分析手段8は、前記入力音声5を分析して、音声のスペクトル包絡情報である線形予測パラメータを抽出する。次いで線形予測パラメータ符号化手段9が抽出した前記線形予測パラメータを量子化し、それに対応する符号を多重化手段3に出力するとともに、量子化した線形予測パラメータを適応音源探索手段31、第2の目標音声生成手段33、駆動音源探索手段35に出力する。

【0054】ピッチ分析手段25は入力音声5を分析してピッチ周期Pを抽出する。次いで遅延パラメータ探索範囲決定手段26は前記ピッチ周期Pより、適応音源ベクトルを探索する際の遅延パラメータ1の探索範囲 $l_{min} \leq 1 \leq l_{max}$ を、例えば式(1)に従って決定し、入力音声アップサンプリング手段27、音源信号アップサンプリング手段29、適応音源探索手段31に出力する。ここで、 ΔP は例えば $P/10$ とする。

$$\begin{aligned} l_{min} &= P - \Delta P \\ l_{max} &= P + \Delta P \end{aligned} \quad (1)$$

【0055】入力音声アップサンプリング手段27は遅延パラメータ探索範囲決定手段26から入力された遅延パラメータの探索範囲に応じたサンプリングレートに入力音声5を、例えば音源信号を符号化する単位であるフレーム区間でアップサンプリングし、目標音声生成手段28に出力する。ここで、アップサンプリングレートは、例えば以下のように決定する。

$l_{min} < 45$ の場合、4倍にアップサンプリングする。
 $45 \leq l_{min} < 65$ の場合、2倍にアップサンプリングする。
 $65 \leq l_{min}$ の場合、アップサンプリングしない。

【0056】目標音声生成手段28は前記入力音声アップサンプリング手段27から入力されたアップサンプリングしたフレーム長の入力音声を、適応音源探索手段31から入力される遅延パラメータ1に対して、例えば周期1毎に分割し、この分割した遅延パラメータ1に対応したベクトル長毎の入力音声を加算平均することによ

り、遅延パラメータ1に対応したベクトル長の目標音声ベクトルを生成し、適応音源探索手段31と第2の目標音声生成手段33に出力する。ここで、遅延パラメータ1は整数値の他、非整数有理数もととり、1の存在範囲に応じて、例えば l_{int} を整数値遅延とした場合に以下の値をとり得るとする。

$1 < 45$ の場合、 l_{int} 、 $l_{int} + 1/4$ 、 $l_{int} + 1/2$ 、 $l_{int} + 3/4$

$45 \leq 1 < 65$ の場合、 l_{int} 、 $l_{int} + 1/2$

$65 \leq 1$ の場合、 l_{int}

図2にフレーム長の入力音声から生成する遅延パラメータ1に対応したベクトル長の目標音声ベクトルの例を示す。ここで、 $1 \geq$ フレーム長の場合は前記加算平均を行わず、フレーム長の入力音声を目標音声ベクトルとする。

【0057】音源信号アップサンプリング手段29は、音源信号生成手段15から入力される過去に生成した音源信号を、前記遅延パラメータ探索範囲決定手段26から入力された前記遅延パラメータの探索範囲に応じた適応音源探索に必要な区間のみを、前記遅延パラメータの探索範囲に応じたサンプリングレートにアップサンプリングし、適応音源符号帳30に出力する。ここで、アップサンプリングレートは、例えば以下のように決定する。

$1 < 45$ の区間、4倍にアップサンプリングする。

$45 \leq 1 < 65$ の区間、2倍にアップサンプリングする。

$65 \leq 1$ の区間、アップサンプリングしない。

【0058】適応音源符号帳30は前記音源信号アップサンプリング手段29から入力されたアップサンプリングした音源信号から、適応音源探索手段31より入力される遅延パラメータ1に対応したベクトル長の適応音源ベクトルを適応音源探索手段31に出力する。ここで、前記適応音源ベクトルは遅延パラメータ1に対して過去1サンプルの音源信号を切り出したものであり、 $1 \geq$ フレーム長の場合は1サンプル過去からフレーム長の音源信号を切り出したものとする。

【0059】適応音源探索手段31は、合成フィルタを備え、線形予測パラメータ符号化手段9から入力される量子化した線形予測パラメータを用いて合成フィルタのインパルス応答を求める。次いで、 $l_{min} \leq 1 \leq l_{max}$ の範囲の遅延パラメータ1に対して、前記適応音源符号帳30から入力される適応音源ベクトルを前記インパルス応答を用いて反復的に計算して合成し、合成音声ベクトルを生成する。そして、前記目標音声生成手段28から入力された目標音声ベクトルと前記合成音声ベクトルとの聴覚重み付き歪みを求める。次に、前記歪みを比較評価し、前記歪みが最小になる遅延パラメータLとそれに対応する適応音源利得 β を求め、前記遅延パラメータ

Lと適応音源利得 β の符号を多重化手段3と駆動音源符号帳34に出力するとともに、前記遅延パラメータLに対応する適応音源ベクトルに前記適応音源利得 β を乗じた適応音源信号を生成し、フレーム音源生成手段32と第2の目標音声生成手段33に出力する。ここで、前記適応音源信号は、 $L < \text{フレーム長}$ の場合はLサンプル、 $L \geq \text{フレーム長}$ の場合はフレーム長の信号である。

【0060】フレーム音源生成手段32は前記適応音源探索手段31から入力された適応音源信号から、例えば周期L毎に繰り返して周期化して、フレーム長の適応音源信号を生成し、音源信号生成手段15に出力する。

【0061】第2の目標音声生成手段33は前記適応音源探索手段31から入力された適応音源信号を線形予測パラメータ符号化手段9から入力される量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、目標音声生成手段28から入力された目標音声ベクトルと前記合成音声ベクトルとの差分を求め、これを第2の目標音声ベクトルとして駆動音源探索手段35に出力する。

【0062】駆動音源符号帳34には、例えばランダム雑音から生成したN個の駆動音源ベクトルが記憶されており、駆動音源探索手段35より入力される駆動音源符号iに対応した駆動音源ベクトルを前記遅延パラメータLに対応したベクトル長で切り出し、出力する。ここで、 $L \geq \text{フレーム長}$ の場合は、フレーム長の駆動音源ベクトルを出力する。

【0063】駆動音源探索手段35はN個の駆動音源ベクトルに対して、前記駆動音源符号帳34から入力される切り出した駆動音源ベクトルを前記線形予測パラメータ符号化手段9から入力される量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、前記第2の目標音声生成手段33から入力された第2の目標音声ベクトルと前記合成音声ベクトルとの聴覚重み付き歪みを求める。次に、前記歪みを比較評価し、前記歪みが最小になる駆動音源符号Iとそれに対応する駆動音源利得 γ を求め、前記駆動音源符号Iと駆動音源利得 γ の符号を多重化手段3に出力するとともに、前記駆動音源符号Iに対応する駆動音源ベクトルに前記駆動音源利得 γ を乗じた駆動音源信号を生成し、第2のフレーム音源生成手段36に出力する。

【0064】第2のフレーム音源生成手段36は前記駆動音源探索手段35から入力された駆動音源信号から、例えば周期L毎に繰り返して周期化して、フレーム長の駆動音源信号を生成し、音源信号生成手段15に出力する。

【0065】音源信号生成手段15は前記フレーム音源生成手段32から入力されたフレーム長の適応音源信号と前記第2のフレーム音源生成手段36から入力されたフレーム長の駆動音源信号とを加算して音源信号を生成し、音源信号アップサンプリング手段29に出力する。

【0066】以上符号化が終了した後、多重化手段3は前記量子化した線形予測パラメータに対応する符号、遅延パラメータL、駆動音源信号I、及び音源利得 β 、 γ に対応する符号を伝送路6に送出する。以上がこの実施の形態1の音声符号化装置に特徴的な動作である。

【0067】つづいて、復号化部2について説明する。まず多重化手段3の出力を受けた分離手段4は、線形予測パラメータの符号→線形予測パラメータ復号化手段16

遅延パラメータL→適応音源復号化手段39、駆動音源符号帳41

音源利得 β の符号→適応音源復号化手段39

駆動音源符号I、音源利得 γ の符号→駆動音源復号化手段42

にそれぞれ出力する。

【0068】適応音源復号化手段39は、まず前記遅延パラメータLを音源信号アップサンプリング手段37と適応音源符号帳38に出力する。音源信号アップサンプリング手段37は、音源信号生成手段21から入力される過去に生成した音源信号を、前記適応音源復号化手段39から入力された前記遅延パラメータLの値に応じた適応音源ベクトル生成に必要な区間のみを、前記遅延パラメータLの値に応じたサンプリングレートにアップサンプリングし、適応音源符号帳38に出力する。ここで、アップサンプリングレートは符号化部における音源信号アップサンプリング手段29と同様に決定する。

【0069】適応音源符号帳38は前記音源信号アップサンプリング手段37から入力されたアップサンプリングした音源信号から、適応音源復号化手段39より入力される遅延パラメータLに対応したベクトル長の適応音源ベクトルを適応音源復号化手段39に出力する。ここで、前記適応音源ベクトルは遅延パラメータLに対して過去Lサンプルの音源信号を切り出したものであり、 $L \geq \text{フレーム長}$ の場合はLサンプル過去からフレーム長の音源信号を切り出したものとする。

【0070】適応音源復号化手段39は、前記適応音源利得 β の符号から適応音源利得 β を復号化し、前記適応音源符号帳38から入力された適応音源ベクトルに前記適応音源利得 β を乗じた適応音源信号を生成し、フレーム音源生成手段40に出力する。フレーム音源生成手段40は前記適応音源復号化手段39から入力された適応音源信号から、例えば周期L毎に繰り返して周期化して、フレーム長の適応音源信号を生成し、音源信号生成手段21に出力する。

【0071】駆動音源符号帳41は符号化側の駆動音源符号帳34と同じN個の駆動音源ベクトルが記憶されており、駆動音源復号化手段42より入力される駆動音源符号Iに対応した駆動音源ベクトルを前記遅延パラメータLに対応したベクトル長で切り出し、駆動音源復号化手段42に出力する。

【0072】駆動音源復号化手段42は、前記駆動音源利得 γ の符号から駆動音源利得 γ を復号化し、前記駆動音源符号帳41から入力された切り出した駆動音源ベクトルに前記駆動音源利得 γ を乗じた駆動音源信号を生成し、第2のフレーム音源生成手段43に出力する。第2のフレーム音源生成手段43は前記駆動音源復号化手段42から入力された駆動音源信号から、例えば周期 L 毎に繰り返して周期化して、フレーム長の駆動音源信号を生成し、音源信号生成手段21に出力する。

【0073】音源信号生成手段21は前記フレーム音源生成手段40から入力されたフレーム長の適応音源信号と前記第2のフレーム音源生成手段43から入力されたフレーム長の駆動音源信号とを加算して音源信号を生成し、音源信号アップサンプリング手段37と合成フィルタ22に出力する。合成フィルタ22は前記音源信号生成手段21から入力された音源信号を、線形予測パラメータ復号化手段16から入力される線形予測パラメータを用いて線形予測合成し、出力音声7を出力する。以上がこの実施の形態1の音声復号化装置に特徴的な動作である。

【0074】この実施の形態1によれば、最適な遅延パラメータを決定する際に、遅延パラメータ1がフレーム長より短い場合は入力音声を周期的に加算平均してベクトル長1の目標音声ベクトルを生成し、これに対してベクトル長1の適応音源ベクトルを線形予測合成して生成した合成音声ベクトルとの歪みを評価し、また、最適な駆動音源符号を決定する際にも、ベクトル長1の駆動音源ベクトルを線形予測合成して生成した合成音声ベクトルを歪み評価に用いることにより、合成音声の品質の劣化を回避し、少ない演算量で品質の良い合成音声を生成することができる。

【0075】実施の形態2。上記実施の形態1では、フレーム音源生成手段32、40及び第2のフレーム音源生成手段36、43において、遅延パラメータ L に対応したベクトル長の適応音源信号あるいは駆動音源信号を周期 L 毎に繰り返して周期化し、フレーム長の適応音源信号あるいは駆動音源信号を生成しているが、前記遅延パラメータ L に対応したベクトル長の適応音源信号あるいは駆動音源信号を、例えば周期 L 毎に波形補間するとして、フレーム間で補間し、フレーム長の適応音源信号あるいは駆動音源信号を生成するとしてもよい。

【0076】この実施の形態2によれば、フレーム間の音源信号の変化が滑らかになり、合成音声の再現性を良好にし、品質を向上させることができる。

【0077】実施の形態3。上記実施の形態1、2では、遅延パラメータ L に対応したベクトル長の適応音源信号と駆動音源信号から、フレーム音源生成手段及び第2のフレーム音源生成手段を用いてフレーム長の適応音源信号とフレーム長の駆動音源信号を生成し、これらを加算してフレーム長の音源信号を生成しているが、遅延

パラメータ L に対応したベクトル長の適応音源信号と駆動音源信号とを加算して遅延パラメータ L に対応したベクトル長の音源信号を生成し、これを例えば周期 L 毎に繰り返して周期化して、フレーム長の音源信号を生成するとしてもよい。

【0078】実施の形態4。上記実施の形態1では、符号化部、復号化部ともに新たな構成をとっているが、符号化部を実施の形態1の符号化部とし、復号化部を図12に記載の従来の復号化部としてもよい。

【0079】実施の形態5。上記実施の形態1では、目標音声生成手段28においてフレーム長の入力音声から遅延パラメータ1に対応したベクトル長の目標音声ベクトルを生成しているが、図3に示すように、遅延パラメータ1に対応したベクトル長の整数倍長の入力音声から目標音声ベクトルを生成するとしてもよい。

【0080】この実施の形態5によれば、目標音声ベクトルを生成する際の平均化処理において、ベクトル長が異なるベクトルを扱う必要がなく、簡易に処理することができる。また、フレーム長を越えた入力音声を音声符号化する際の評価に用いることにより、該フレームの合成音声に該フレーム以後に与える影響も加味して符号を決定することになり、合成音声の再現性を良好にし、品質を向上させることができる。

【0081】実施の形態6。上記実施の形態1では、目標音声生成手段28において入力音声から遅延パラメータ1に対応したベクトル長の目標音声ベクトルを生成する際に単純平均しているが、図4に示すように、各遅延パラメータ1に対応したベクトル長の入力音声のパワーに応じた重みで、例えば、パワーが大きいほど重みを大きくするとして、加重平均するとしてもよい。

【0082】この実施の形態6によれば、目標音声ベクトルを生成する際の平均化処理において、入力音声のパワーが大きい部分により重みをつけて音声符号化することにより、主観品質に与える影響が大きい合成音声のパワーの大きな部分の再現性が良好になり、品質を向上させることができる。

【0083】実施の形態7。上記実施の形態1では、目標音声生成手段28において入力音声から遅延パラメータ1に対応したベクトル長の目標音声ベクトルを生成する際に単純平均しているが、図5に示すように、各遅延パラメータ1に対応したベクトル長の入力音声間の相互相関値に応じた重みで、例えば、他の各遅延パラメータ1に対応したベクトル長の入力音声との相関が低い場合は重みを小さくするとして、加重平均するとしてもよい。

【0084】この実施の形態7によれば、目標音声ベクトルを生成する際の平均化処理において、入力音声の周期1の周期性をもつとした場合に相関が低い部分の重みを小さくして音声符号化することにより、ピッチ周期が変動している入力音声に対しても1ピッチ周期に対応し

た歪みが小さい目標音声ベクトルを生成することができ、合成音声の再現性を良好にし、品質を向上させることができる。

【0085】実施の形態 8. 上記実施の形態 1 では、目標音声生成手段 28 において入力音声から遅延パラメータ 1 に対応したベクトル長の目標音声ベクトルを生成する際に単純平均しているが、図 6 に示すように、各遅延パラメータ 1 に対応したベクトル長の入力音声間の位置に応じた重みで、例えば、フレーム境界近傍の入力音声に対して重みを大きくすると、加重平均するとしてもよい。

【0086】この実施の形態 8 によれば、目標音声ベクトルを生成する際の平均化処理において、フレーム境界近傍の入力音声に重みを大きくして目標音声ベクトルを生成し、符号化することにより、フレーム境界近傍の合成音声の再現性を良好にすることができ、フレーム間の合成音声の変化を滑らかにすることができる。この効果は、実施の形態 2 における音源信号をフレーム間で補間して生成する場合に、特に顕著となる。

【0087】実施の形態 9. 上記実施の形態 1 では、目標音声生成手段 28 において入力音声から遅延パラメータ 1 に対応したベクトル長の目標音声ベクトルを生成する際に入力音声を周期 1 毎に加算平均しているが、図 7 に示すように、入力音声を切り出す位置を、例えば各遅延パラメータ 1 に対応したベクトル長の入力音声間の相互相関が最大になるように、微調整し、加算平均するとしてもよい。

【0088】この実施の形態 9 によれば、目標音声ベクトルを生成する際の平均化処理において、遅延パラメータ 1 に対応したベクトル長の入力音声間の相互相関が大きくなるように切り出し位置を微調整することにより、ピッチ周期が変動している入力音声に対しても 1 ピッチ周期に対応した歪みが小さい目標音声ベクトルを生成することができ、合成音声の再現性を良好にし、品質を向上させることができる。

【0089】実施の形態 10. 図 8 は、この発明の実施の形態 10 である音声符号化装置及び音声復号化装置の全体構成を示すブロック図である。この図において図 1 と同一の部分には同一の符号を付してあるので、説明を省略する。

【0090】図 8 において、図 1 と比べて新たな構成は次のとおりである。44 は入力音声をアップサンプリングする入力音声アップサンプリング手段、45 はピッチ周期に対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段、46、51 はピッチ周期に対応したベクトル長の駆動音源ベクトルを出力する駆動音源符号帳、47 は駆動音源ベクトルから得られる合成音声ベクトルの目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する駆動音源探索手段、48 は第 2 の駆動音源ベクトル探索におけるピッ

チ周期に対応したベクトル長の目標音声ベクトルを生成する第 2 の目標音声生成手段、49、54 はピッチ周期に対応したベクトル長の第 2 の駆動音源ベクトルを出力する第 2 の駆動音源符号帳、50 は第 2 の駆動音源ベクトルから得られる合成音声ベクトルの第 2 の目標音声ベクトルに対する歪みを評価し、歪みが最小となる駆動音源ベクトルを探索する第 2 の駆動音源探索手段、52 はピッチ周期に対応したベクトル長の駆動音源信号を復号化する駆動音源復号化手段、53 はピッチ周期に対応したベクトル長の駆動音源信号からフレーム長の駆動音源信号を生成するフレーム音源生成手段、55 はピッチ周期に対応したベクトル長の第 2 の駆動音源信号を復号化する第 2 の駆動音源復号化手段、56 はピッチ周期に対応したベクトル長の第 2 の駆動音源信号からフレーム長の駆動音源信号を生成する第 2 のフレーム音源生成手段である。

【0091】以下、上記の新たな構成を中心に動作を説明する。まず、符号化部 1 において、ピッチ分析手段 25 は入力音声 5 を分析してピッチ周期 P を抽出し、多重化手段 3、入力音声アップサンプリング手段 44、目標音声生成手段 45、駆動音源符号帳 46、第 2 の駆動音源符号帳 49 に出力する。ここで、ピッチ周期 P は整数値の他、非整数有理数もと、 P の存在範囲に応じて、例えば P_{int} を整数値ピッチ周期とした場合に以下の値をとり得るとする。

$P < 4.5$ の場合、 P_{int} 、 $P_{int} + 1/4$ 、 $P_{int} + 1/2$ 、 $P_{int} + 3/4$

$4.5 \leq P < 6.5$ の場合、 P_{int} 、 $P_{int} + 1/2$

$6.5 \leq P$ の場合、 P_{int}

【0092】入力音声アップサンプリング手段 44 はピッチ分析手段 25 から入力されたピッチ周期に応じたサンプリングレートに入力音声 5 を、例えば音源信号を符号化する単位であるフレーム区間でアップサンプリングし、目標音声生成手段 45 に出力する。ここで、アップサンプリングレートは、例えば以下のように決定する。

$P < 4.5$ の場合、4 倍にアップサンプリングする。

$4.5 \leq P \leq 6.5$ の場合、2 倍にアップサンプリングする。

$6.5 \leq P$ の場合、アップサンプリングしない。

【0093】目標音声生成手段 45 は前記入力音声アップサンプリング手段 44 から入力されたアップサンプリングしたフレーム長の入力音声を、ピッチ分析手段 25 から入力されるピッチ周期 P に対応して、例えば周期 P 毎に加算平均することにより、ベクトル長 P の目標音声ベクトルを生成し、駆動音源探索手段 47 と第 2 の目標音声生成手段 48 に出力する。ここで、 $P \geq$ フレーム長の場合は前記加算平均を行わず、フレーム長の入力音声为目标音声ベクトルとする。

【0094】駆動音源符号帳 46 には、例えばランダム

雑音から生成したN個の駆動音源ベクトルが記憶されており、駆動音源探索手段47より入力される駆動音源符号 i に対応した駆動音源ベクトルを前記ピッチ分析手段25より入力されたピッチ周期 P に対応したベクトル長で切り出し、出力する。ここで、 $P \geq$ フレーム長の場合、フレーム長の駆動音源ベクトルを出力する。

【0095】駆動音源探索手段47はN個の駆動音源ベクトルに対して、前記駆動音源符号帳46から入力される切り出した駆動音源ベクトルを線形予測パラメータ符号化手段9から入力される量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、前記目標音声生成手段45から入力された目標音声ベクトルと前記合成音声ベクトルとの聴覚重み付き歪みを求める。次に、前記歪みを比較評価し、前記歪みが最小になる駆動音源符号 I とそれに対応する駆動音源利得 γ を求め、前記駆動音源符号 I と駆動音源利得 γ の符号を多重化手段3に出力するとともに、前記駆動音源符号 I に対応する駆動音源ベクトルに前記駆動音源利得 γ を乗じた駆動音源信号を生成し、第2の目標音声生成手段48に出力する。

【0096】第2の目標音声生成手段48は前記駆動音源探索手段47から入力された駆動音源信号を線形予測パラメータ符号化手段9から入力される量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、前記目標音声生成手段45から入力された目標音声ベクトルと前記合成音声ベクトルとの差分を求め、これを第2の目標音声ベクトルとして第2の駆動音源探索手段50に出力する。

【0097】第2の駆動音源符号帳49には、例えばランダム雑音から生成したN個の駆動音源ベクトルが記憶されており、第2の駆動音源探索手段50より入力される駆動音源符号 j に対応した第2の駆動音源ベクトルを前記ピッチ分析手段25から入力されたピッチ周期 P に対応したベクトル長で切り出し、出力する。ここで、 $P \geq$ フレーム長の場合は、フレーム長の駆動音源ベクトルを出力する。

【0098】第2の駆動音源探索手段50はN個の駆動音源ベクトルに対して、前記第2の駆動音源符号帳49から入力される切り出した第2の駆動音源ベクトルを前記線形予測パラメータ符号化手段9から入力される量子化した線形予測パラメータを用いて線形予測合成して合成音声ベクトルを生成する。そして、前記第2の目標音声生成手段48から入力された第2の目標音声ベクトルと前記合成音声ベクトルとの聴覚重み付き歪みを求める。次に、前記歪みを比較評価し、前記歪みが最小になる第2の駆動音源符号 J とそれに対応する第2の駆動音源利得 γ_j を求め、前記第2の駆動音源符号 J と第2の駆動音源利得 γ_j の符号を多重化手段3に出力する。

【0099】以上符号化が終了した後、多重化手段3は前記量子化した線形予測パラメータに対応する符号、ピ

ッチ周期 P 、駆動音源符号 I 、 J 、及び音源利得 γ 、 γ_j に対応する符号を伝送路6に送出する。以上がこの実施の形態10の音声符号化装置に特徴的な動作である。

【0100】つづいて、復号化部2について説明する。まず多重化手段3の出力を受けた分離手段4は、線形予測パラメータの符号→線形予測パラメータ復号化手段16

ピッチ周期 P →駆動音源符号帳51、第2の駆動音源符号帳54

駆動音源符号 I 、音源利得 γ の符号→駆動音源復号化手段52

第2の駆動音源符号 J 、音源利得 γ_j の符号→第2の駆動音源復号化手段55にそれぞれ出力する。

【0101】駆動音源符号帳51は符号化側の駆動音源符号帳46と同じN個の駆動音源ベクトルが記憶されており、駆動音源復号化手段52より入力される駆動音源符号 I に対応した駆動音源ベクトルを前記ピッチ周期 P に対応したベクトル長で切り出し、駆動音源復号化手段52に出力する。

【0102】駆動音源復号化手段52は、前記駆動音源利得 γ の符号から駆動音源利得 γ を復号化し、前記駆動音源符号帳51から入力された切り出した駆動音源ベクトルに前記駆動音源利得 γ を乗じた駆動音源信号を生成し、フレーム音源生成手段53に出力する。フレーム音源生成手段53は前記駆動音源復号化手段52から入力された駆動音源信号から、例えば周期 P 毎に繰り返して周期化して、フレーム長の駆動音源信号を生成し、音源信号生成手段21に出力する。

【0103】第2の駆動音源符号帳54は符号化側の第2の駆動音源符号帳49と同じN個の駆動音源ベクトルが記憶されており、第2の駆動音源復号化手段55より入力される第2の駆動音源符号 J に対応した第2の駆動音源ベクトルを前記ピッチ周期 P に対応したベクトル長で切り出し、第2の駆動音源復号化手段55に出力する。

【0104】第2の駆動音源復号化手段55は、前記第2の駆動音源利得 γ_j の符号から駆動音源利得 γ_j を復号化し、前記第2の駆動音源符号帳54から入力された切り出した第2の駆動音源ベクトルに前記駆動音源利得 γ_j を乗じた第2の駆動音源信号を生成し、第2のフレーム音源生成手段56に出力する。第2のフレーム音源生成手段56は前記第2の駆動音源復号化手段55から入力された第2の駆動音源信号から、例えば周期 P 毎に繰り返して周期化して、第2のフレーム長の駆動音源信号を生成し、音源信号生成手段21に出力する。

【0105】音源信号生成手段21は前記フレーム音源生成手段53から入力されたフレーム長の駆動音源信号と前記第2のフレーム音源生成手段56から入力された第2のフレーム長の駆動音源信号とを加算して音源信号を生成し、合成フィルタ22に出力する。合成フィルタ

22は前記音源信号生成手段21から入力された音源信号を、線形予測パラメータ復号化手段16から入力される線形予測パラメータを用いて線形予測合成し、出力音声7を出力する。以上がこの実施の形態10の音声復号化装置に特徴的な動作である。

【0106】この実施の形態10によれば、入力音声のピッチ周期Pがフレーム長より短い場合は入力音声を周期的に加算平均してベクトル長Pの目標音声ベクトルを生成し、これに対してベクトル長Pの駆動音源ベクトルを線形予測合成して生成した合成音声ベクトルとの歪みを評価することにより、合成音声の品質の劣化を回避し、少ない演算量で品質の良い合成音声を生成することができる。

【0107】

【発明の効果】以上詳述したように、請求項1～請求項4、請求項6、請求項9、請求項14、請求項16～請求項23に記載の発明によれば、音声符号化装置に、入力音声から遅延パラメータに対応したベクトル長の目標音声ベクトルを生成する目標音声生成手段と、過去に生成した音源信号から前記遅延パラメータに対応したベクトル長の適応音源ベクトルを生成する適応音源符号帳と、前記適応音源ベクトルから得られる合成音声ベクトルの前記目標音声ベクトルに対する歪みを評価し、歪みが最小となる適応音源ベクトルを探索する適応音源探索手段と、前記歪みが最小となる適応音源ベクトルからフレーム長の音源信号を生成するフレーム音源生成手段とを備えているので、合成音声の品質の劣化を回避し、少ない演算量で品質の良い合成音声を生成することができる。

【0108】また、請求項5に記載の発明によれば、目標音声ベクトルのベクトル長が有理数をとるようにしたので、入力音声から目標音声ベクトルを生成する際に、入力音声のサンプリング周期に因らず精度良く目標音声ベクトルを生成することができ、合成音声の品質の劣化を回避し、少ない演算量で品質の良い合成音声を生成することができる。

【0109】また、請求項7に記載の発明によれば、目標音声生成手段は遅延パラメータに対応したベクトル長の整数倍長の入力音声をベクトル長毎に分割し、前記ベクトル長毎の入力音声を加重平均して目標音声ベクトルを生成するものとしたので、目標音声ベクトルを生成する際の平均化処理において、ベクトル長が異なるベクトルを扱う必要がなく、簡易に処理することができ、合成音声の品質の劣化を回避し、少ない演算量で品質の良い合成音声を生成することができる。

【0110】また、請求項8に記載の発明によれば、目標音声ベクトルを生成するベクトル長の整数倍長の入力音声をフレーム長以上とするものとしたので、フレーム長を越えた入力音声を音声符号化する際の評価に用いることにより、該フレームの合成音声は該フレーム以後に

与える影響も加味して符号を決定することになり、合成音声の再現性を良好にし、品質を向上させることができる。

【0111】また、請求項10に記載の発明によれば、ベクトル長毎の入力音声に関わる特徴量は少なくとも入力音声のパワー情報を含むようにしたので、入力音声のパワーが大きい部分により重みをつけて音声符号化することにより、主観品質に与える影響が大きい合成音声のパワーの大きな部分の再現性が良好になり、品質を向上させることができる。

【0112】また、請求項11に記載の発明によれば、ベクトル長毎の入力音声に関わる特徴量は少なくとも入力音声の相関情報を含むようにしたので、入力音声の周期1の周期性をもつとした場合に相関が低い部分の重みを小さくして音声符号化することにより、ピッチ周期が変動している入力音声に対しても1ピッチ周期に対応した歪みが小さい目標音声ベクトルを生成することができ、合成音声の再現性を良好にし、品質を向上させることができる。

【0113】また、請求項12に記載の発明によれば、目標音声生成手段はベクトル長毎の入力音声の時間関係に応じて、入力音声を前記ベクトル長毎に加重平均して目標音声ベクトルを生成する際の重みを決定するようにしたので、フレーム境界近傍の入力音声に重みを大きくして目標音声ベクトルを生成し、符号化することにより、フレーム境界近傍の合成音声の再現性を良好にすることができ、フレーム間の合成音声の変化を滑らかにすることができる。

【0114】また、請求項13に記載の発明によれば、目標音声生成手段は入力音声をベクトル長毎に加重平均する際、前記ベクトル長毎の入力音声の時間関係を微調整するようにしたので、ベクトル長1の入力音声間の相互関係が大きくなるように切り出し位置を微調整することにより、ピッチ周期が変動している入力音声に対しても1ピッチ周期に対応した歪みが小さい目標音声ベクトルを生成することができ、合成音声の再現性を良好にし、品質を向上させることができる。

【0115】また、請求項15に記載の発明によれば、フレーム音源生成手段はベクトル長の音源ベクトルをフレーム間で補間して音源信号を生成するようにしたので、フレーム間の音源信号の変化が滑らかになり、合成音声の再現性を良好にし、品質を向上させることができる。

【図面の簡単な説明】

【図1】 この発明の実施の形態1である音声符号化装置及び音声復号化装置の全体構成を示すブロック図である。

【図2】 この発明の実施の形態1における目標音声生成手段の動作の一例を示す説明図である。

【図3】 この発明の実施の形態5における目標音声生

成手段の動作の一例を示す説明図である。

【図 4】 この発明の実施の形態 6 における目標音声生成手段の動作の一例を示す説明図である。

【図 5】 この発明の実施の形態 7 における目標音声生成手段の動作の一例を示す説明図である。

【図 6】 この発明の実施の形態 8 における目標音声生成手段の動作の一例を示す説明図である。

【図 7】 この発明の実施の形態 9 における目標音声生成手段の動作の一例を示す説明図である。

【図 8】 この発明の実施の形態 10 である音声符号化装置及び音声復号化装置の全体構成を示すブロック図である。

【図 9】 従来の音声符号化復号化装置の一例の全体構成を示すブロック図である。

【図 10】 従来の音声符号化復号化装置における適応音源ベクトルの一例を示す説明図である。

【図 11】 従来の改良された音声符号化復号化装置における適応音源ベクトルの一例を示す説明図である。

【図 12】 従来の音声符号化復号化装置の他の異なる一例の全体構成を示すブロック図である。

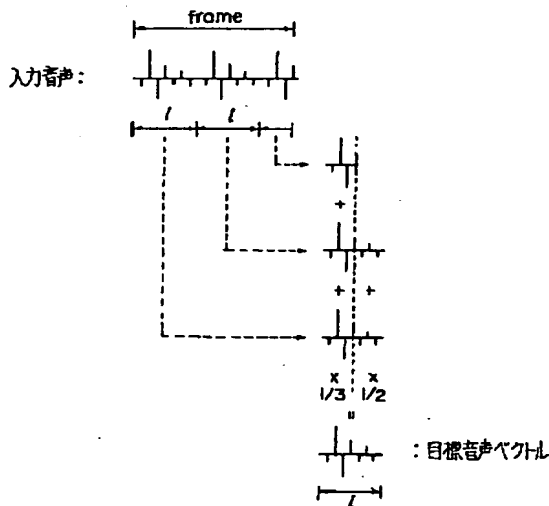
【図 13】 従来の音声符号化復号化装置における周期化した駆動音源ベクトルの一例を示す説明図である。

【符号の説明】

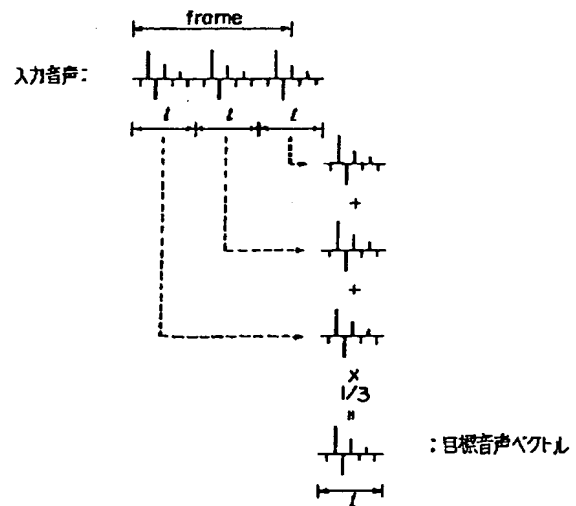
1 符号化部、2 復号化部、3 多重化手段、4 分

離手段、5 入力音声、6 伝送路、7 出力音声、8 線形予測パラメータ分析手段、9 線形予測パラメータ符号化手段、10、17 適応音源符号帳、11 適応音源探索手段、12 誤差信号生成手段、13、19 駆動音源符号帳、14 駆動音源探索手段、15、21 音源信号生成手段、16 線形予測パラメータ復号化手段、18 適応音源復号化手段、20 駆動音源復号化手段、22 合成フィルタ、23、24 駆動音源符号帳、25 ピッチ分析手段、26 遅延パラメータ探索範囲決定手段、27 入力音声アップサンプリング手段、28 目標音声生成手段、29 37 音源信号アップサンプリング手段、30、38 適応音源符号帳、31 適応音源探索手段、32、40 フレーム音源生成手段、33 第 2 の目標音声生成手段、34、41 駆動音源符号帳、35 駆動音源探索手段、36、43 第 2 のフレーム音源生成手段、39 適応音源復号化手段、42 駆動音源復号化手段、44 入力音声アップサンプリング手段、45 目標音声生成手段、46、51 駆動音源符号帳、47 駆動音源探索手段、48 第 2 の目標音声生成手段、49、54 第 2 の駆動音源符号帳、50 第 2 の駆動音源探索手段、52 駆動音源復号化手段、53 フレーム音源生成手段、55 第 2 の駆動音源復号化手段、56 第 2 のフレーム音源生成手段。

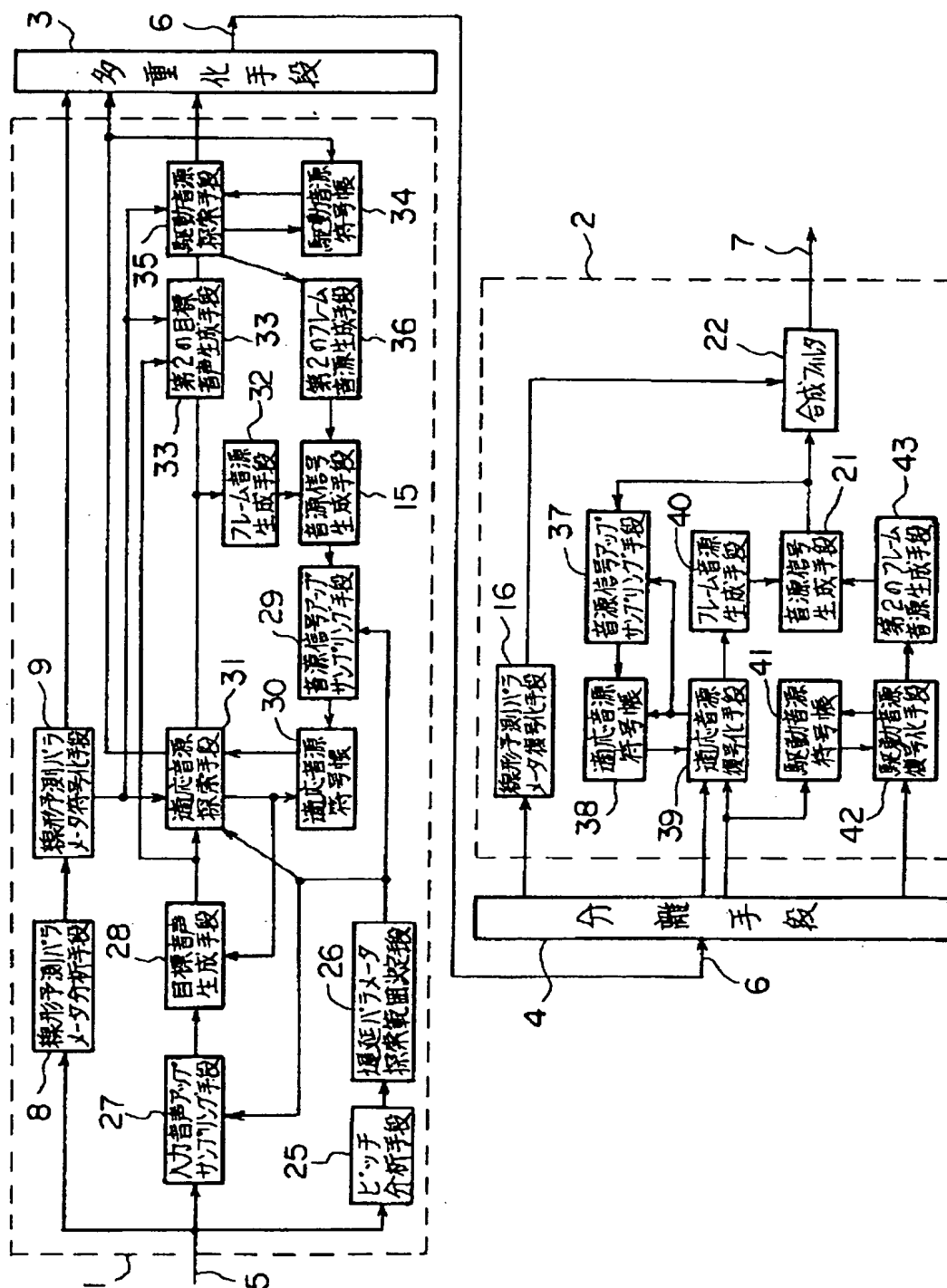
【図 2】



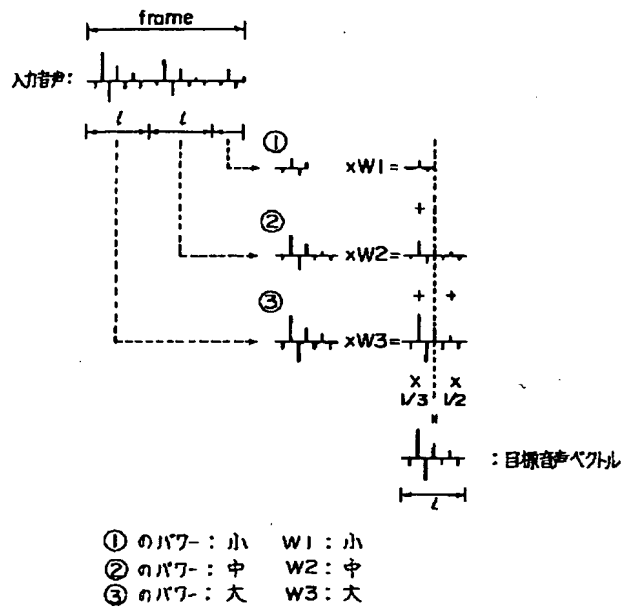
【図 3】



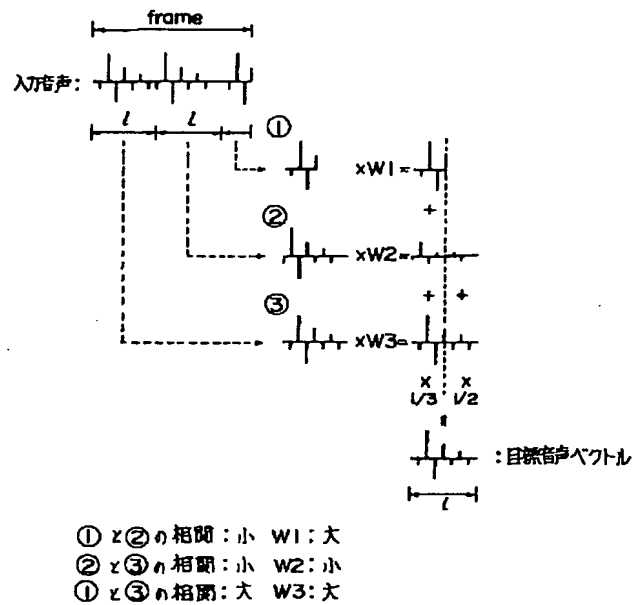
【図 1】



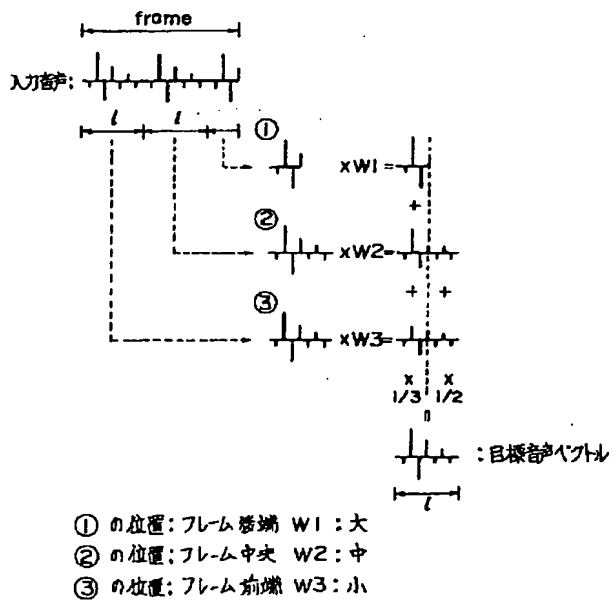
【図 4】



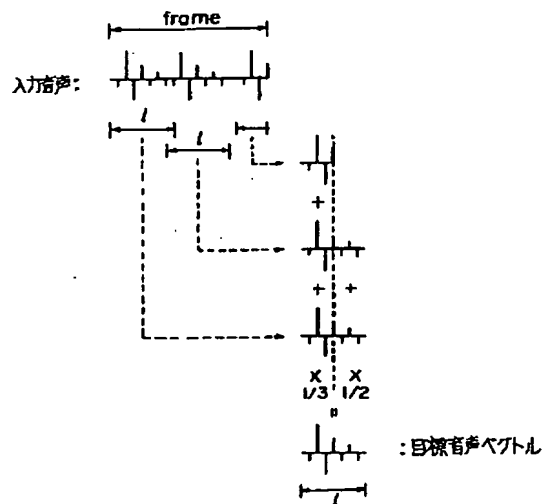
【図 5】

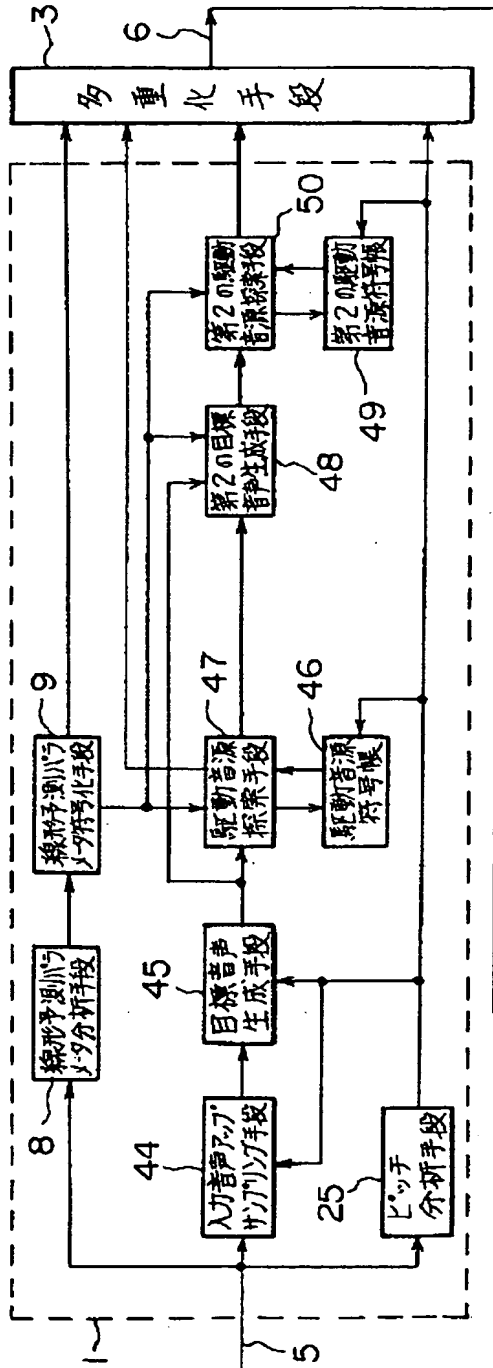


【図 6】

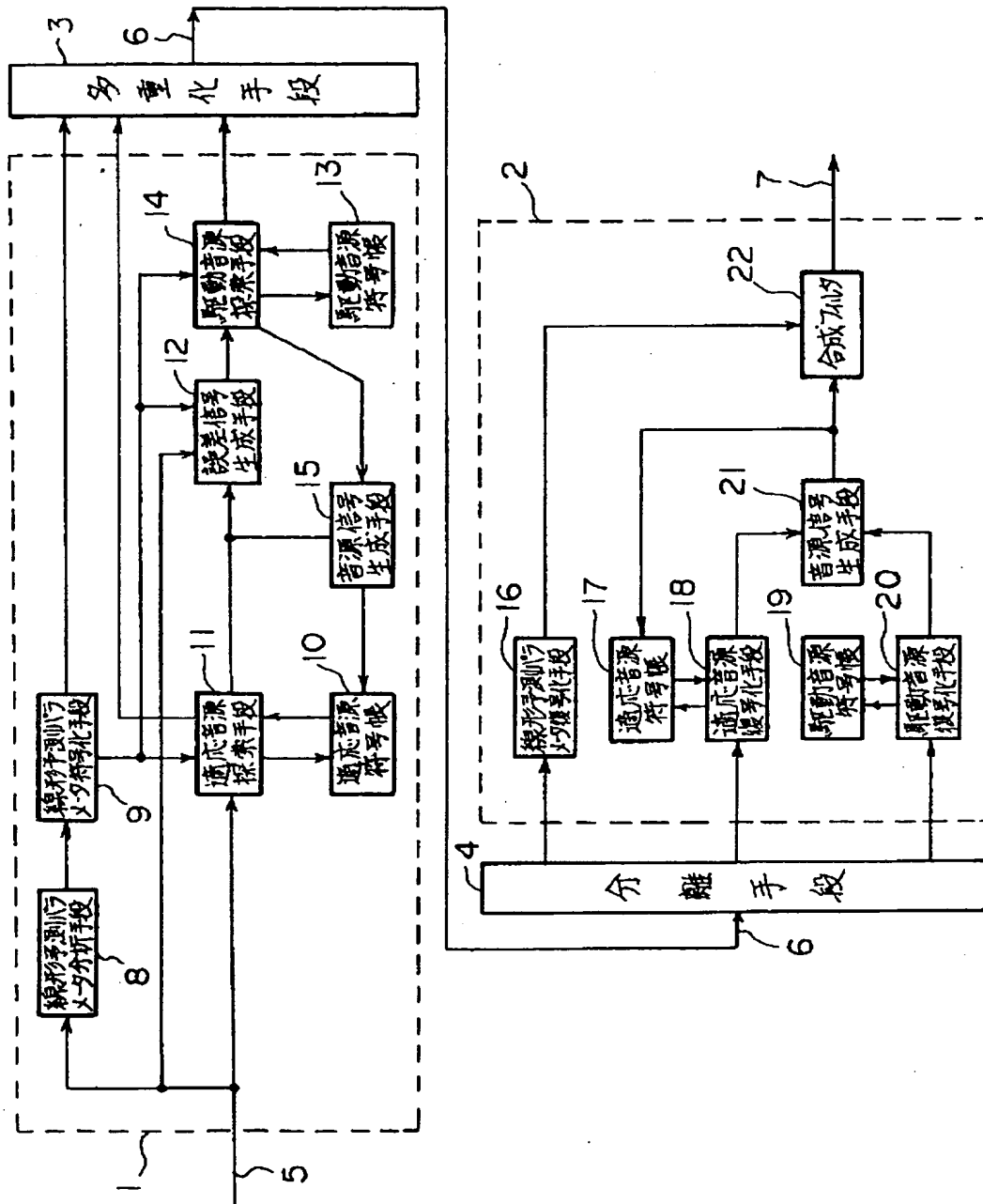


【図 7】

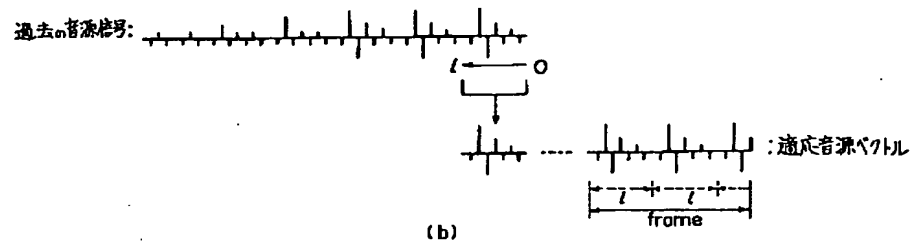
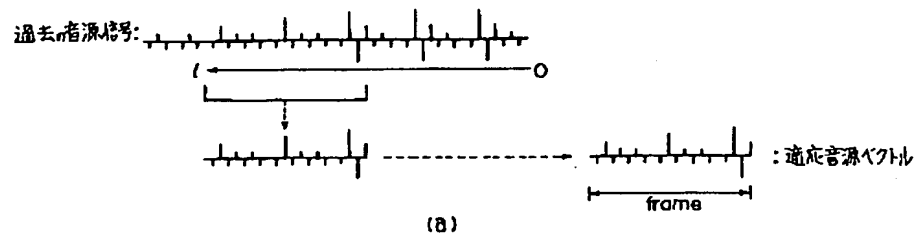




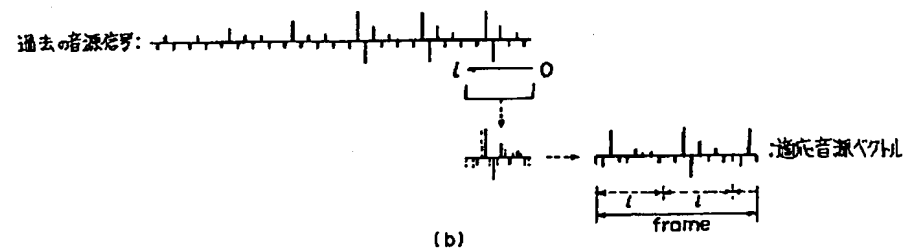
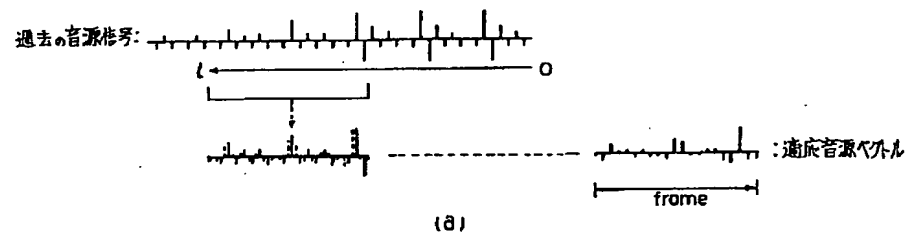
【図 9】



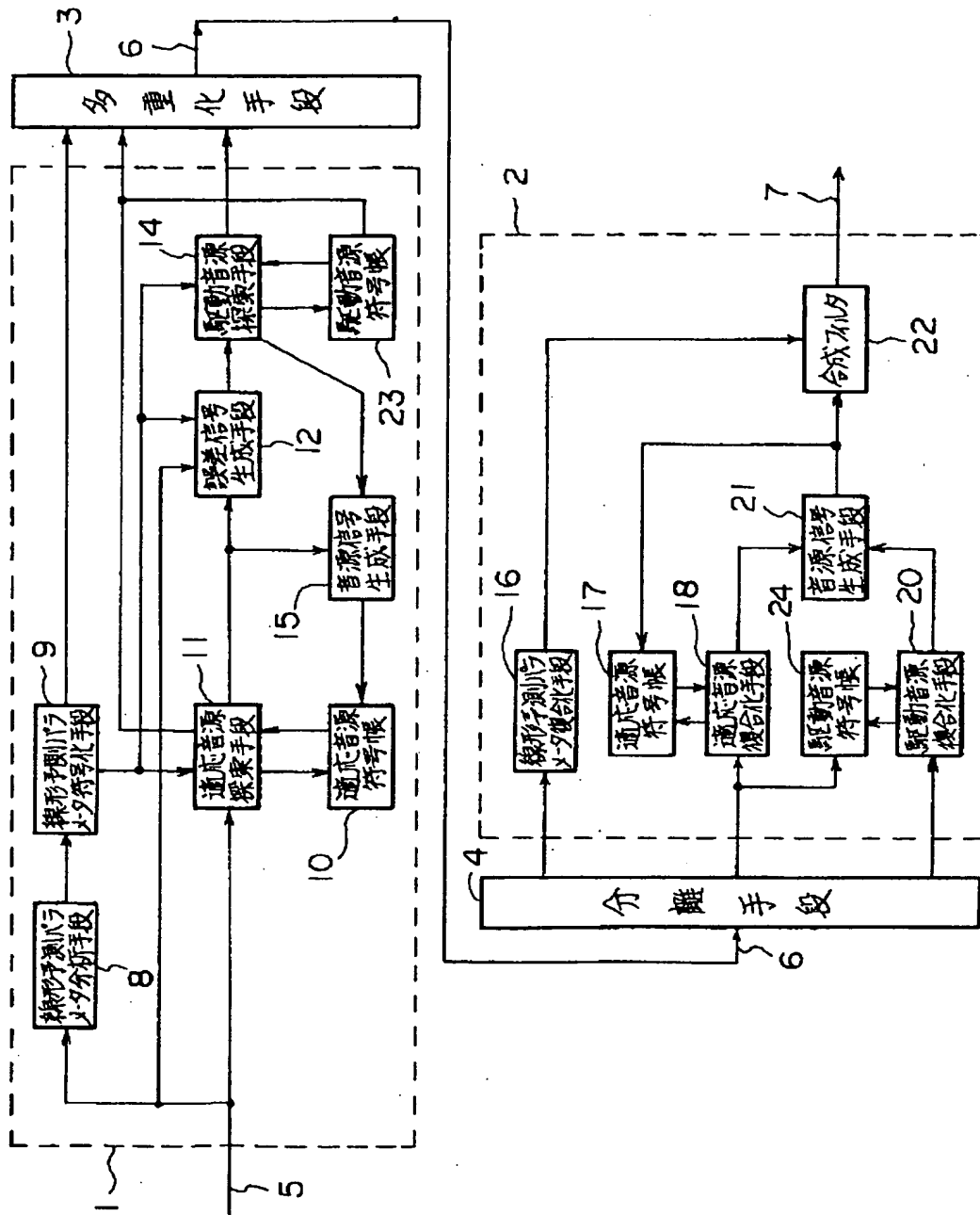
【図 10】



【図 11】



【図12】



【図 13】

